

*T. reesei* **bg/4** nucleotide sequence (non-coding is in bold, coding is in regular font)

[illegible]

BGL4 Beta-Glucosidase and Nucleic Acids  
Encoding the Same  
Dunn-Coleman et al.  
SN# Unassigned  
Docket No. GC696  
Sheet 2 of 4

AGTCTATTGCCAGCTGGCCGCCGGAACCAAACACCGTGTGTCTATGCAGACGGGCAACCCCGAGGAGATGCCTCTGGC  
TCCAGCCACAGCCCGCGGTATCATCAGCGCTTGTATCGGGCAACAGACGGGCAACTCCATTGCCGACGTGTCTTTGGCG  
ACTTAAACCCCTTGGGCAAGCTGTCCCTCAGCTTCCCAAGCGCTGCAGGACAACCCCGGTGTCTTCAACTTCCGCAACCG  
AGGCGGGCGCAGCTGTATCGCGAGGACGTTTACGTCGGGTACAGGTACTACAGTTTCCGACAGACGCTCAATTTC  
CTTTGGCCACGGCCTGTCTTAAACCACTATTCCTTCCCTTCCAACTCTCCGTGTCTCAAGACGCGAAGCTGAGCGGTGT  
CCCTTCCCGTGAAGACACACCGGCTCCGTGCCGGGCGACAGTGGCCCAAGCTTACGTCAAGCCCTCAACGCGGCCAAGA  
TTTAAACCGCCCGCTCAAGAGACTCAAGGGCTTCGAAAGGTGCACTGCACCCGCGAGACAGGCGGTGACAACTCGAG  
ACAGGAGAAAGTACGTCTGGCTATTTTGTATGAGGACGCGGACTCAGTGTGTGTCTGAAAGGGTGACTATGAGTTTATCG  
TGAGACACAGCAGCGCAGCGAGGATTGGGTTTGGCTTACGGGGTAAAGTTTACGGGGAGAGACGATTTGGTGTGGTCTTGGCG  
TTGAAAGTCTGGTGCATCATCTTTGGCAGATTGAAATCCAGTCACTTGTGCATATGCGGCCGATGAATGAGAACCAACGATCAT  
TGTCTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA

FOBT2T"00022001

Figure 2-A

predicted *T. reesei* BGL4 amino acid sequence

MADIDVEAIL	10	20	30	40	50	60	70
KLTLAEKVD							
LLAGIDFWHT							
KALPKHGVP							
LRFTDGNV							
RGTKFFNGVP							
AACFCGCTSL							
GSTFNQTLLE	80	90	100	110	120	130	140
EAGKRWGKEA							
IAKSAHVILG							
PTINMQRSPL							
GGRGFESIGE							
DPFLAGLCAA							
ALIRGIQSTG							
VQATIKHFLC	150	160	170	180	190	200	210
NDQEDRRMV							
QSVITERALR							
EIYALPFOIA							
VRDSQGFAM							
TAYNGINGVS							
CSENPKYLDG							
MLRKEWGWDG	220	230	240	250	260	270	280
LIMSDWYGT							
STTEAVVAGL							
DLEMPGPPRF							
RGETLKENVS							
NGKPFHVID							
QFAREVILQFV							
KKCAASGVTE	290	300	310	320	330	340	350
NGPETTVNNT							
PETAALLRKV							
GNEGIVLLKN							
ENNVLPSSKK							
KKTILIVGPNA							
KQATYHGGGS							
AALRAYYAVT	360	370	380	390	400	410	420
PFDGLSKQLE							
TPPSYTVGAY							
TTVPPIILGEQ							
CLTPDGAGFM							
RWRVNEPPG							
TPNRQHIDEL							
FFTKTDMHLV	430	440	450	460	470	480	490
DYYHPKAAVT							
WYADMEGTYT							
ADEDCTYELG							
LIVVCGTAKAY							
VDDQLVVDNA							
TKQVPGDAFF							
GSATREETGR	500	510	520	530	540	550	560
INLVKGNTYK							
FKIEFGSAPT							
YTLKGDITVP							
GHGSLRVGGC							
KVIDDQAEIE							
KSVALAKEHD							
QVIICAGLNA	570	580	590	600	610	620	630
DNETEGADRA							
SMKLPGVLDQ							
LIADVAAANP							
NTVVVMQTGT							
PEEMPWLDAT							
PAVIQAWYGG							

108121\*00022001

Figure 2-B

Predicted *T. reesei* BGL4 amino acid sequence

NETGNSIADV	VFGDYNPSGK	LSLSFFKRLQ	DNPAFLNFR	EAGRTLYGED	VVVGYYEYEF	ADKDVNFPFG	
640	650	660	670	680	690	700	
HGLSYTTFAF	SNLSVSHKDG	KLVSLSVKN	TGSVPGAQVA	QIYVPLQAA	KINRPVKELK	GFAKVELOPG	
710	720	730	740	750	760	770	
ETKAVTIEEQ	EKIYAAYFDE	ERDQWCVEKG	DYEVIVSDSS	AAKDGVALRG	KFTVGETYWW	SGV	
780	790	800	810	820	830	833	